The Friendship Paradox for Social Networks

Frank den Hollander
Leiden University, The Netherlands

Seminar Series, 20 April 2026, Bengaluru, India,
Centre for Networked Intelligence (CNI), Indian Institute of Science (IISc),
Department of Electrical, Electronics, and Computer Sciences (EECS).

SUMMARY

I. The friendship paradox.

II. Quantification via convergence.

III. Two examples.

I. THE FRIENDSHIP PARADOX

§ THE FRIENDSHIP PARADOX IN WORDS

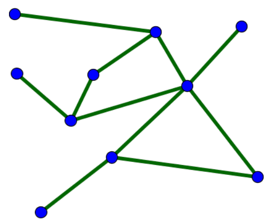
"On Average, Our Friends Have More Friends Than We Do!"

Feld, 1991

a social network

§ SOCIAL NETWORKS MODELLED AS GRAPHS

• For $n \in \mathbb{N}$, let G_n be a finite undirected graph with n vertices labeled by $[n] = \{1, \ldots, n\}$. Let d_i be the degree of vertex i.



• Let $\Delta_{i,n}$ be the friendship-bias of vertex i, which is defined as the difference between the average degree of the neighbours of i and the degree of i itself, i.e.,

$$\Delta_{i,n} = \left[\frac{\sum_{j\in[n]} A_{ij} d_j}{d_i} - d_i\right] \mathbb{1}_{\{d_i \neq 0\}}, \qquad i \in [n],$$

where A is the adjacency matrix of G_n, i.e., A_{ij} = number of edges between $i \neq j$ and A_{ii} = twice number of self-loops at i.

§ THE FRIENDSHIP PARADOX IN A FORMULA

For any graph G_n , the average friendship-bias is non-negative, i.e.,

$$\Delta_{[n]} = \frac{1}{n} \sum_{i \in [n]} \Delta_{i,n} \geq 0$$

Equality holds if and only if all the connected components of G_n are regular, i.e., have constant degree.

PROOF: Write, using the relation $d_i = \sum_{j\in[n],d_j\neq 0} A_{ij}$,

$$\Delta_{[n]} = \frac{1}{n}\sum_{\substack{i\in[n]\\ d_i\neq 0}}\left(\sum_{j\in[n]}\frac{A_{ij}d_j}{d_i} - d_i\right) = \frac{1}{n}\sum_{\substack{i\in[n]\\ d_i\neq 0}}\sum_{\substack{j\in[n]\\ d_j\neq 0}} A_{ij}\left(\frac{d_j}{d_i} - 1\right)$$

$$= \frac{1}{2n}\sum_{\substack{i\in[n]\\ d_i\neq 0}}\sum_{\substack{j\in[n]\\ d_j\neq 0}} A_{ij}\left(\frac{d_j}{d_i} - 1 + \frac{d_i}{d_j} - 1\right)$$

$$= \frac{1}{2n}\sum_{\substack{i\in[n]\\ d_i\neq 0}}\sum_{\substack{j\in[n]\\ d_j\neq 0}} A_{ij}\left(\sqrt{\frac{d_j}{d_i}} - \sqrt{\frac{d_i}{d_j}}\right)^2 \geq 0.$$

Equality holds if and only if $i \mapsto d_i$ is constant on each connected component. $\square$

§ APPLICATIONS

The friendship paradox plays a role in:

- Epidemics: vaccination strategies.
- Psychology: perception of inequality.
- Politics: voting prediction via polling.
- Communication: spread of viral content.
- Statistics: sampling bias in surveys.
- Mathematics: geometry of networks.

§ KEY QUESTION

Can we quantify the friendship paradox and analyse the quantification for key examples of large random graphs?

Why large random graphs? Because social networks are typically vast in size and complex in shape. The key object to look at is the friendship-bias empirical distribution

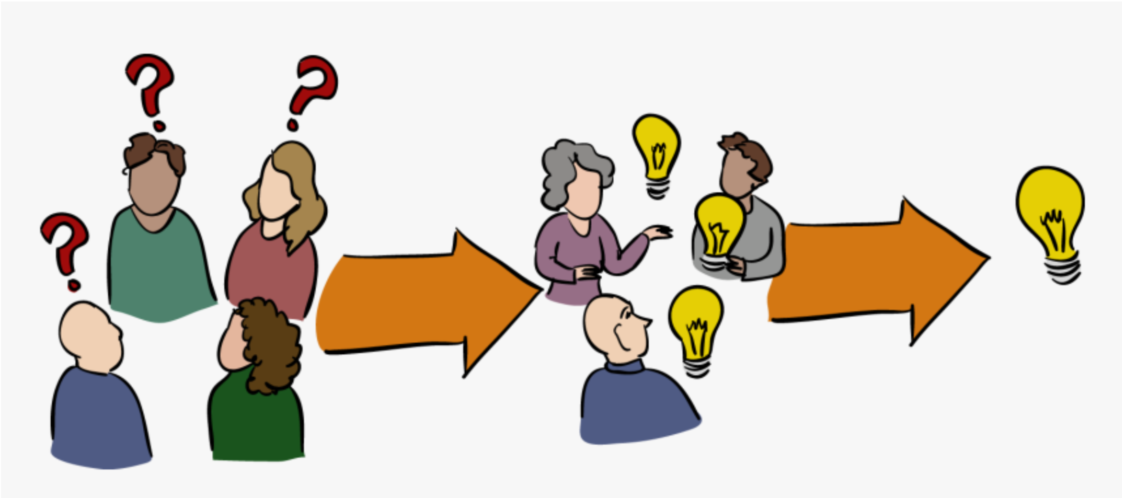
$$\mu_n = \frac{1}{n} \sum_{i \in [n]} \delta_{\Delta_{i,n}}.$$

Pal, Yu, Novick, Swami, Amotz, Bar-Noy 2019

Cantwell, Kirkley, Newman 2021

Hazra, den Hollander, Parvaneh 2025

II. QUANTIFICATION VIA CONVERGENCE



§ LOCAL CONVERGENCE

In what follows we look at a sequence of random graphs $(G_n)_{n\in\mathbb{N}}$. We need the following notion of convergence, where $\mathbb{P}_n$ denotes the probability law of G_n.

DEFINITION

$(G_n)_{n\in\mathbb{N}}$ is said to converge locally in probability to the pair (G_∞, ϕ) when G_n viewed relative to a randomly drawn vertex converges in law, in an appropriate topology as $n \to \infty$ under the law $\mathbb{P}_n$, to a random graph G_∞ with root vertex ϕ.

THEOREM 1 Hazra, den Hollander, Parvaneh 2025

If $(G_n)_{n\in\mathbb{N}}$ *converges locally in probability to* (G_∞, ϕ)*, then* μ_n *converges weakly to* μ *in probability, where* μ *is the law of the friendship-bias of* ϕ *in* G_∞*.*

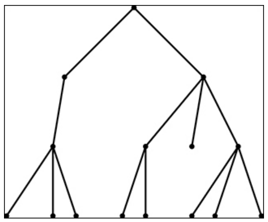
Theorem 1 says that

$$\lim_{n \to \infty} \mu_n(B) = \mu(B) \text{ in probability for all } B \subset \mathbb{R} \text{ Borel.}$$

Thus, we have a full characterisation of the limit law μ in terms of the friendship-bias of the root in the limiting rooted random graph.

The above characterisation allows us to fully quantify the friendship-paradox!

§ ROOTED RANDOM TREES



Many sparse random graphs ($\approx$ bounded degrees) are locally tree-like, which represents a structural simplification.

Let (G_∞, ϕ) be an almost surely locally finite rooted random tree. Let d_ϕ be the degree of ϕ, and let d_j be the degree of neighbour j of ϕ. Then

$$\Delta_\phi = \left[\frac{1}{d_\phi}\sum_{j=1}^{d_\phi} d_j - d_\phi\right]\mathbb{1}_{\{d_\phi \neq 0\}},$$

is the friendship-bias of ϕ in G_∞ and μ is the law of Δ_ϕ.

§ SIGNIFICANCE

DEFINITION Hazra, den Hollander, Parvaneh 2025

The friendship paradox is called

$$\text{significant} \quad \text{when} \quad \mu([0,\infty)) \geq \tfrac{1}{2},$$

$$\text{insignificant} \quad \text{when} \quad \mu([0,\infty)) < \tfrac{1}{2}.$$

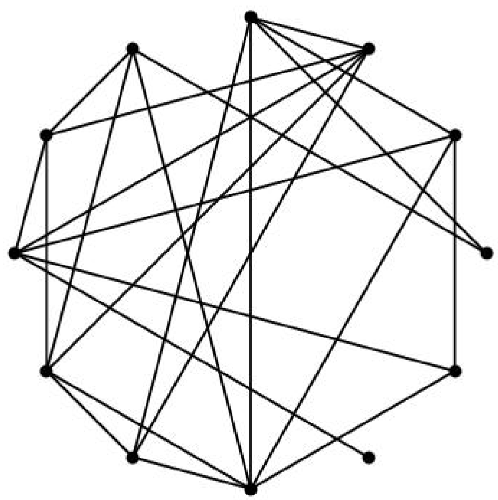
We already know that $\int_{\mathbb{R}} x\mu(\mathrm{d}x) \geq 0$. Significance means that the vertices with a non-negative friendship-bias are at least as numerous as the vertices with a negative friendship-bias, i.e., the median of μ is non-negative.

III. TWO EXAMPLES

R. van der Hofstad, *Random Graphs and Complex Networks*,
Cambridge Series in Statistical and Probabilistic Mathematics,
Volume 1 (2017), Volume 2 (2024).

§ HOMOGENEOUS ERDŐS-RÉNYI RANDOM GRAPH

For $\lambda \in (0, \infty)$ and $n \in \mathbb{N}$, let $\mathrm{ER}_n(\lambda)$ be the random graph in which each pair of distinct vertices in $i, j \in [n]$ is independently connected by an edge with probability $\frac{\lambda}{n} \wedge 1$.



a realisation of $\mathrm{ER}_n(\lambda)$ with $n = 12$ and $\lambda = 6$

BASIC FACT:

As $n \to \infty$, $\mathrm{ER}_n(\lambda)$ converges locally to a rooted random tree with offspring distribution $\mathrm{Poisson}(\lambda)$.

THEOREM 2a

$$\lim_{\lambda \downarrow 0} \mathbb{E}_\mu[\Delta_\phi] = 0, \quad \lim_{\lambda \downarrow 0} \mathbb{E}_\mu[\Delta_\phi^2] = 0,$$

$$\lim_{\lambda \to \infty} \mathbb{E}_\mu[\Delta_\phi] = 1, \quad \lim_{\lambda \to \infty} \mathbb{E}_\mu[\Delta_\phi^2] = \infty.$$

THEOREM 2b

For every $\lambda \in (0, \infty)$,

$$\mathbb{P}_\mu\{\Delta_\phi \geq x\} \sim \frac{\lambda\, \mathrm{e}^{-2\lambda}}{\sqrt{2\pi x}} \exp\left\{ -x \log\left(\frac{x}{\lambda \mathrm{e}}\right)\right\}, \quad x \to \infty.$$

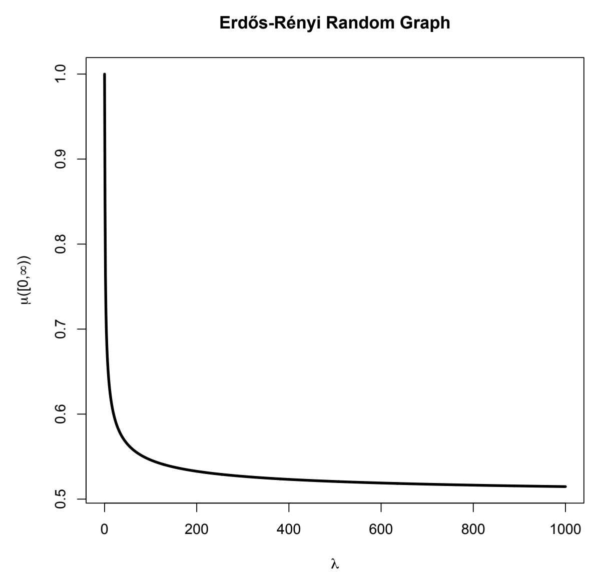
THEOREM 2c

(a) *For every* $\lambda \in (0,\infty)$,

$$\mu([0,\infty)) = \sum_{k\in\mathbb{N}_0} \frac{e^{-\lambda}\lambda^k}{k!} \sum_{l\geq k(k-1)} \frac{e^{-\lambda k}(\lambda k)^l}{l!}.$$

(b)

$$\lim_{\lambda\downarrow 0} \mu([0,\infty)) = 1, \qquad \lim_{\lambda\to\infty} \mu([0,\infty)) = \tfrac{1}{2}.$$

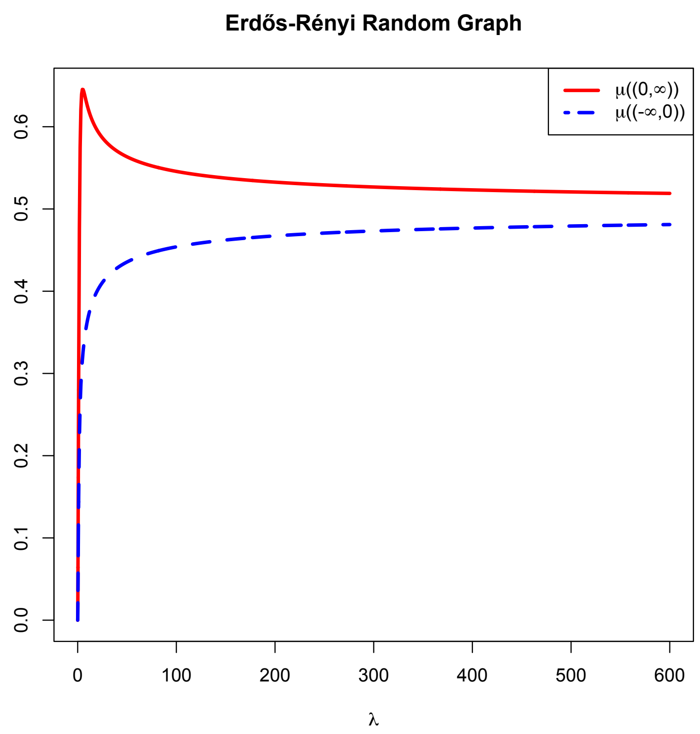


Refined numerical plot:

$\mu((0,\infty))$ (= positive vertices),

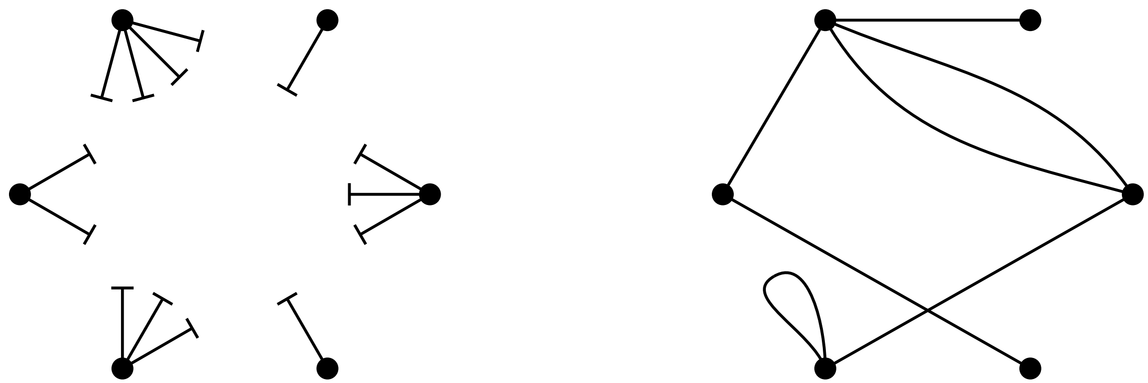
$\mu((-\infty,0))$ (= negative vertices).

The former outweigh the latter, indicating a strong form of significance.



§ CONFIGURATION MODEL

Let $\mathrm{CM}_n(d_n)$ be the random graph drawn uniformly at random from the set of all graphs with a prescribed degree sequence d_n. Such a graph can be generated via a simple algorithm.



$n = 6$

prescribed degree sequence $d_6 = (1, 3, 1, 3, 2, 4)$

randomly pair half-edges

BASIC FACT:

Let D_n be the degree of a uniformly chosen vertex. Suppose that, for some D with $\mathbb{P}(D > 0) = 1$ and $\mathbb{E}[D] < \infty$,

$$D_n \to D \text{ in distribution}, \quad \mathbb{E}[D_n] \to \mathbb{E}[D], \qquad n \to \infty.$$

Then $\mathrm{CM}_n(d_n)$ converges locally to a random tree with

- offspring distribution $p = (p_k)_{k \in \mathbb{N}_0}$ at the root ϕ,
- offspring distribution $p^\star = (p^\star_k)_{k \in \mathbb{N}_0}$ at all other vertices,

where

$$p_k = \mathbb{P}\{D = k\}, \qquad p^\star_k = \frac{(k+1)p_{k+1}}{\mathbb{E}[D]}.$$

SPECIAL CASE: the zeta offspring distribution

$$p_k = \frac{1}{\zeta(\tau)} k^{-\tau}, \quad k \in \mathbb{N}, \qquad \tau \in (2, \infty).$$

THEOREM 3a

For general offspring distribution,

$$\mathbb{E}_\mu[\Delta_\phi] = \frac{\mathrm{Var}(D)}{\mathbb{E}[D]}.$$

If $\mathbb{E}[D^2] < \infty$*, then also* $\mathbb{E}_\mu[\Delta_\phi^2]$ *is computable. In particular,*

$$\mathrm{Var}_\mu(\Delta_\phi) \geq \mathrm{Var}(D).$$

THEOREM 3b

For the zeta offspring distribution,

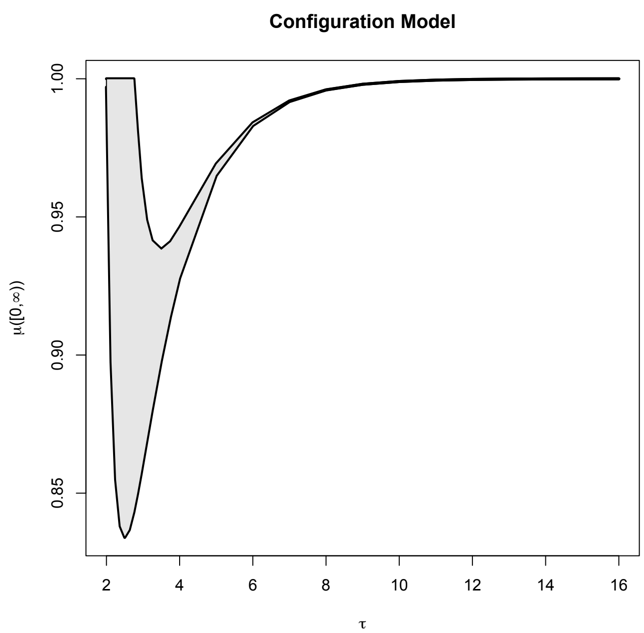
$$\mathbb{P}_\mu\{\Delta_\phi \geq x\} \asymp x^{-(\tau-2)}, \qquad x \to \infty.$$

THEOREM 3c

For the zeta offspring distribution,

$$\mu([0,\infty)) > \tfrac{1}{2} \quad \forall \tau \in (2,\infty),$$

$$\mu([0,\infty)) \to 1, \quad \tau \downarrow 2, \qquad \mu([0,\infty)) \to 1, \quad \tau \to \infty.$$

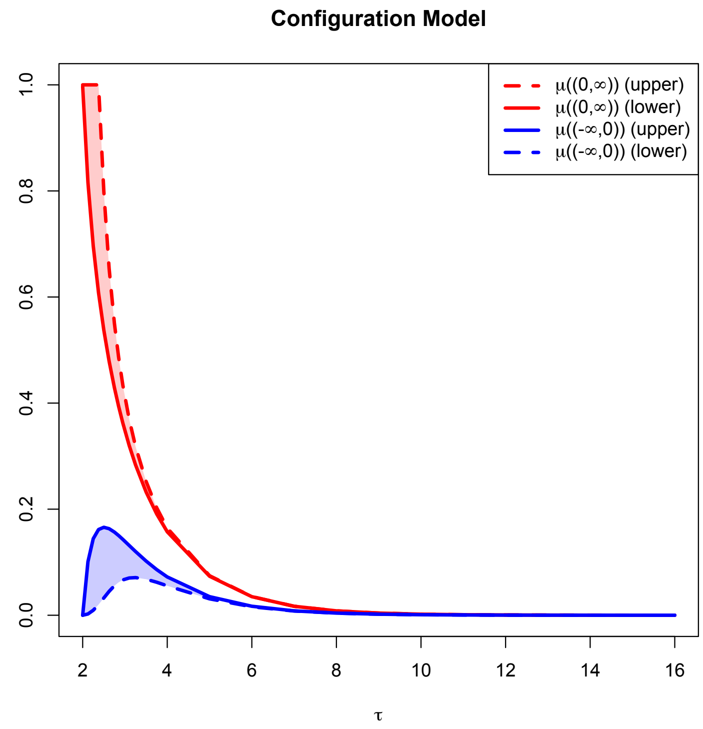


Refined numerical plot:

$\mu((0,\infty))$ (= positive vertices),

$\mu((-\infty,0))$(= negative vertices).

The former outweigh the latter, indicating a strong form of significance.




THANK YOU